

# R Basics for Data Science Theory

This documents briefly covers the R Basics concepts for data science.

## 1. Printing

- This command print means "show on the screen".

## 2. Variables

- Variables are storage containers for data.
- For example, a food variable could contain "Spaghetti" or "Rice" or any other value.
- The value of a variable can change, depending on conditions or on information passed to the program.
- Naming conventions for r
  - A variable can be a combination of letters, digits, period (.) and underscore (\_).
  - A variable must start with a letter or a period. If it starts with a period, it cannot be followed by a digit.
  - Reserved words in R cannot be used as variables.
- Every programming language has reserved words that cannot be used in variable names.
- Once we create variables we are then able to perform operations taking into account of the stored values.

## 3. Comments

- Comments are text notes that are added to the program with the aim of explaining how your program works.
- Comments don't have any effect on your program.
- Comments have the following goals:
  - To explain what a particular code does.
  - To explain something which might not be obvious to the reader.
  - To clarify your intention behind a certain line or block of code.
  - To serve as a reminder to change something in the future.

## 4. Arithmetic Operators

- Arithmetic operators are used to perform mathematical operations such as addition, subtraction, multiplication, division, modulus, exponentiation, etc.
- The addition operator will produce the sum of numeric operands or string concatenation.
- The subtraction operator will subtract two operands and produce their difference.

- The division operator will produce the quotient of its operands where the left operand is the dividend and the right operand is the divisor.
- The multiplication operator will produce the product of the operands.
- The modulus operator will return the remainder left over when one operand is divided by a second operand.
- The exponentiation operator will return the result of raising the first operand to the power second operand.
- The increment operator will increment (add one to) its operand and return a value.
- The decrement operator will decrement (subtract one from) its operand and return a value.

## 5. Data Types

- A data type is a classification that allows specifying which type of value a variable contains.
- It also allows us to determine what type of mathematical operations can be applied to any given data.
- The integer value is a positive or negative whole number.
- The float value is a real number with floating-point representation.
- The string value is a sequence of characters.

## 6. Lists

- A list is a compound data type that allows for the storage of items of different data types.
- Lists are mutable, which means that we can change an item in the lists by accessing it directly.
- Each element or value that is inside a list is called an item.
- We work with lists when we would like to work with many related values i.e. food, songs, emails, etc.
- The various operations that we can perform on lists are indexing (which allows us to access its items), slicing, updating, appending, removing.

## 7. Conditional (if) Statements

- Conditional if statements are used for decision making. They will run code if only when if the statement is true.
- An "if statement" is written by using the if keyword and is used when we need to print out the result when one of the conditions is true or false.
- The "else if statement" is used when we have a third possibility as the outcome. We can use multiple elif conditions to check for the 4th, 5th, 6th possibilities in our code.
- We can also nest our if statements by including if statements within our if statements.

- While working with if statements we also have to take into consideration of curly braces.

## 8. For Loops

- We use a for loop to iterate over a sequence i.e. list, vector, string, etc.
- The loop continues until the last item in the sequence has been reached.
- The body of the for loop is also separated from the rest of the code using indentation.
- We can also have for loops within for loops. This is what we would call nesting.
- We can use a break statement in a for loop to stop the loop before it has looped through all the items.

## 9. Vectors

- Vector is a basic data structure in R that contains elements of the same type.
- The data types of vectors can be logical, integer, double, character, complex or raw.
- There are various ways to create vectors but one of the most common is the concatenation operator.
- This takes arguments and places them all in a vector. Example
  - `x <- c(1, 5, 4, 9, 0)`
- To get the element 5 in the above example, we can use `x[2]`